WPS5709

Policy Research Working Paper 5709

Where on Earth is Everybody?

The Evolution of Global Bilateral Migration 1960–2000

Çağlar Özden
Christopher R. Parsons
Maurice Schiff
Terrie L. Walmsley

The World Bank
Development Research Group
Trade and Integration Team
June 2011

Policy Research Working Paper 5709

Abstract

Global matrices of bilateral migrant stocks spanning 1960–2000 are presented, disaggregated by gender and based primarily on the foreign-born definition of migrants. More than one thousand census and population register records are combined to construct decennial matrices corresponding to the five census rounds between 1960 and 2000. For the first time, a comprehensive picture of bilateral global migration over the second half of the 20th century emerges. The data reveal that the global migrant stock increased from 92 million in 1960 to 165 million in 2000. Quantitatively, migration between developing countries dominates, constituting half of all international migration in 2000. When the partition of India and the dissolution of the Soviet Union are accounted for, migration between developing countries is remarkably stable over the period. Migration from developing to developed countries is the fastest growing component of international migration in both absolute and relative terms. The United States has remained the most important migrant destination in the world, home to one fifth of the world's migrants and the top destination for migrants from some 60 sending countries. Migration to Western Europe has come largely from elsewhere in Europe. The oil-rich Persian Gulf countries emerge as important destinations for migrants from the Middle East and North Africa and South and Southeast Asia. Finally, although the global migrant stock is predominantly male, the proportion of female migrants increased noticeably between 1960 and 2000. The number of women rose in every region except South Asia.

This paper is a product of the Trade and Integration Team, Development Research Group. It is part of a larger effort by the World Bank to provide open access to its research and make a contribution to development policy discussions around the world. Policy Research Working Papers are also posted on the Web at http://econ.worldbank.org. The author may be contacted at cozden@worldbank.org.

Where on Earth is Everybody? The Evolution of Global Bilateral Migration 1960–2000

Çağlar Özden, Christopher R. Parsons, Maurice Schiff, and Terrie L. Walmsley

Keywords: emigration, immigration, labor migration, international migration, gender
JEL Codes: F22, O15, J11, J16

Çağlar Özden (corresponding author; cozden@worldbank.org) is a senior economist in the Development Research Group of the World Bank. Christopher Parsons (lexcrp1@nottingham.ac.uk P) is a consultant at the World Bank and a doctoral candidate at the University of Nottingham. Maurice Schiff (mschiff@worldbank.org) is a lead economist in the Office of the Chief Economist for Latin America at the World Bank. Terrie Walmsley [twalmsle@purdue.edu] is an associate professor and executive director of the Global Trade Analysis Project, Purdue University, and an associate professor at the University of Melbourne. First and foremost the authors thank the United Nations Population Division for spearheading the creation of the Global Migration Database. In particular, they thank Bela Hovy and Hania Zlotnik for their close support and shared vision that ensured the completion of this project. They are grateful to Richard Black, Ronald Skeldon, and especially L. Alan Winters for having the foresight to initiate this project and for their unwavering support. They also extend thanks to the librarians at the British Library, the Library of Congress, and the London School of Economics and to Lorraine Wright at the United States Census Bureau for assistance beyond the call of duty. The authors thank Steven Vertovec and Norbert Winnige of the Max Planck Institute for the Study of Religious and Ethnic Diversity for helping surmount the data issues concerning Germany and the former Soviet Union. They thank Michel Beine, Frédéric Docquier, and the journal editor, as well as three anonymous referees, for advice and comments. They gratefully acknowledge financial support from the World Bank Knowledge for Change Program and Ivar Cederholm's help with administration of the funding. The findings, interpretations, and conclusions expressed in this article are those of the authors and do not necessarily represent the views of the World Bank, its Executive Directors, or the countries they represent.

International migration—the movement of people across national borders—has important economic, social, and political implications. Despite the recent emergence of a dynamic literature, empirical analysis of migration flows and their impact lags behind the policy debate and the theoretical literature. The main reason is the absence of comprehensive and reliable data on international migration patterns and migrant characteristics at either the aggregate or the household level.

The objective of this article is to use data from more than one thousand national censuses and population registers to estimate a complete global origin–destination migration matrix for each decade over 1960–2000. These 226*226 matrices, comprising every country, major territory, and dependency around the world, are divided into periods corresponding to the last five completed census rounds. The gender dimension of international migration over this period is also presented.

The primary source of the raw data is the United Nations Population Division's Global Migration Database, created through the collaboration of the United Nations Population Division, the United Nations Statistics Division, the World Bank, and the University of Sussex (United Nations [2008]). This unique data repository comprises 3,500 individual census and population register records[1] for more than 230 destination countries and territories over the last five decades. The database provides information on international bilateral migrant stocks (by citizenship[2] or place of birth), sex, and age. There is considerable variation, however, in how destination countries collect, record, and disseminate immigration data. Meaningful comparison of destination country records over time is thus often confounded.

In constructing global bilateral migration matrices, several challenges arise. First, destination countries typically classify migrants in different ways—by place of birth, citizenship, duration of stay, or type of visa. Using different criteria for a global dataset generates discrepancies in the data. Second, many geopolitical changes occurred between 1960 and 2000, with many international borders redrawn as new countries emerged and others disappeared. In addition to creating millions of migrants overnight—as when the Soviet Union collapsed—these events complicate the tracking of migrants over time. Third, even when national censuses of destination countries include data on international migrant stocks, the data are presented along aggregate geographic categories rather than by country of origin. Data therefore need to be disaggregated to the country level. Finally, the greatest hurdle is dealing with omitted or missing census data. Very few destination countries—especially developing countries—have conducted rigorous censuses or population registers during every census round over the second half of the twentieth century. Wars, civil strife, lack of funding, and political intransigence are but a few reasons why records may be discontinuous.

[1] Of the 3,500 sources detailed in the overarching UN Global Migration Database, 1,107 were suitable for analysis, once repeated censuses had been removed or combined. Global Migration Database should not be confused with the Trends in International Migrant Stock Database, which lists aggregate migrant stocks for each destination country in the world at five year intervals (United Nations 2006)

[2] The article treats the concepts of nationality and citizenship as analogous and uses the terms interchangeably.

The main contributions of this article lie in identifying and overcoming these challenges in order to construct a consistent and complete set of origin–destination matrices of international migrant stocks for 1960–2000, disaggregated by gender. The starting point is a master set of 226 origin or destination countries and regions. Despite border changes, all migrants are assigned to this master set so that migrations can be meaningfully tracked over time. These assignments, especially in cases where only aggregate data are available, are made using several alternative propensity measures based either on a destination country's propensity to accept international migrants or on an origin country's propensity to send migrants abroad.

Cases of omitted data occur when destination countries do not collect or publicly disseminate the information on migrants. When data from census rounds are missing altogether, the approach taken depends on the extent of the omission (see appendices 3 and 4). When sufficient data are available for other decades, interpolation is used. When not enough data are available, propensity measures are used to generate bilateral data. When a gender breakdown is missing, gender splits are calculated based on supplementary statistics or other data in the matrices (see appendix 5). The resulting migration matrices should be viewed as work in progress, but they are an important step in an ongoing global effort to improve migration data. The matrices can be readily updated as additional or superior information surfaces, and they can easily be extended to include future census rounds.

Bilateral datasets of international migration are rare. Attempts to create them have focused almost exclusively on industrialized countries as destinations because these countries have more accurate and more frequently produced data. Harrison and others (2003) calculate bilateral remittances for the countries of the Organisation for Economic Co-operation and Development (OECD) together with the 27 largest nonmembers. These estimates are based on international bilateral migrant stock data that the authors also provide, although many of the data are derived from the Trends in International Migration (OECD 2002). This report, published annually since 1973, was arguably the most comprehensive guide to international migration for many years and has been the basis for many studies (see, for example, Mayda 2007).

More recently, the OECD has developed a database that provides a comprehensive overview of migration to OECD countries in 2000 (OECD 2008). These data are disaggregated by a number of covariates including age, gender, educational attainment, and place of birth. Another series of papers, again concentrating on the OECD, examines the brain drain in 1990 and 2000 (see, for example, Docquier and Marfouk 2006); migrants' gender (Docquier, Lowell, and Marfouk 2009); age of entry (Beine, Docquier, and Rapoport 2007); and the medical brain drain (Bhargava and Docquier 2007). Parsons and others (2007) construct a matrix encompassing the entire world for the 2000 census round. Until now, this was the most comprehensive global overview of bilateral migrant movements. Ratha and Shaw (2007) use an earlier version of the dataset in a paper focusing on migration between developing countries (generally referred to as South – South migration in the literature) and bilateral remittance flows.

The data in the current article reveal several important patterns. Between 1960 and 2000, the global migrant stock rose from 92 million to 165 million, but fell as a share of world population, from 3.05 percent to 2.71 percent. A large share of the stock in 1960

reflects the partition of India, and in all decades migration within the Soviet Union (and former Soviet Union) accounts for a large proportion of the world migrant stock. A majority of the remaining migrant stocks is due mainly to increasing migration from developing countries to the United States, Western Europe, and the Persian Gulf (referred to as South – North Migration). While the growth in South-North migration has been astonishing, North-North, North-South and South-South migrations all represent declining shares of world migration. Even so, South-South migration dominates global trends numerically. The majority of these migrations are intraregional, within the countries of the former Soviet Union, South Asia, and West Africa. Interregional migrations between developing countries are principally to the Persian Gulf countries.

The United States continues to be the most important destination, home to around one fifth of the world's migrant population and the recipient of the largest migrant flows from no less than 60 countries. At the beginning of the period, most migrants in the United States were born in Europe; today the vast majority comes from Latin America and the Caribbean. This change in the composition of migrant stocks mirrors the wider trend. In 1960, except for migration within the Soviet Union, the majority of migrants were born in Europe and South Asia. In 2000, migration from these regions remained important, but migration from Latin America, East Asia, North Africa, and the Middle East is also prominent. The origin countries most affected by international migration are small, typically island states, mostly in the Pacific or the Caribbean. The destination countries most affected by migration are the countries of the New World (the United States, Canada, Australia, and New Zealand) and the oil-rich Persian Gulf countries.

The data clearly show that international migration is spreading across the globe as migrants widen their destination choices. By 2000, a greater number of migration flows were observed between more country-pairs than at any other time covered in this database. For example, migrants from East Asia and Pacific who once migrated elsewhere within the region now constitute sizable communities across the world. An increasing number of Africans make their homes in Europe and the United States. This diversification is also reflected in destination countries' willingness to accept migrants from ever more diverse backgrounds. This is particularly the case for the United States, Australia, New Zealand, and Canada, all of which select migrants based on qualifications rather than country of origin.

The gender composition of international migration flows has also evolved. Although the global migrant stock is still disproportionately male, the percentage of women in the global migrant stock rose between 1960 and 2000. This increased feminization of international migration is particularly pronounced in the immigrant stocks of Latin America and the Caribbean, Japan, East Asia and Pacific, and Sub-Saharan Africa. These four areas have also experienced the greatest increase in the proportion of female emigrants over the period.

The article is organized as follows. Section I discusses definitions of migrants and how migrants are recorded, describes the raw data, and identifies gaps in knowledge. Section II considers the comparability of migration data and the major challenges in constructing the matrices. It also discusses the conventions and assumptions adopted in meeting the challenges. Given these assumptions, section III investigates the reliability of the estimates, and section IV analyzes the data, highlighting the key patterns in

international migration over 1960–2000. Section V discusses some implications of the study.

I. PRELIMINARIES

Migration data are complex. They almost always come from destination countries, because it is difficult for origin countries to collect demographic data on people who are not living in the country. Unlike trade and financial statistics, which are recorded by both transacting parties, the quality of migration statistics depends almost entirely on the rigor with which destination countries survey the migrants within their borders. In addition, destination countries' recording and dissemination methods can differ greatly. Understanding the analysis in this article requires an understanding of the subtle differences in various sources and definitions, together with an understanding of the inherent inconsistencies between them.[3]

Who Are Classified As Migrants?

The United Nations (1998, p. [6]) defines a migrant as "any person that changes his or her country of usual residence." This broad definition implies a movement from one location to another, the most relevant concept for economic analysis. However, official records apply many different definitions of what constitutes an international migrant. Most common criteria are based on country of birth, country of citizenship, purpose of visit or visa type, place of last permanent residence, and duration of stay.

The two main definitions of migration—being born in or being a citizen of a foreign country—are used most consistently over time and across countries. Citizenship is important for determining an individual's legal rights for employment, voting, and access to public services. The place of birth definition is superior for determining physical movement. Destination countries typically publish migration statistics by either category, mainly according to national migration and citizenship laws. Historically, countries in the Americas and Oceania favor the country of birth definition whereas countries in Asia, Africa, and Europe traditionally adopt a mix of the two definitions.

Individuals may be classified as migrants or nonmigrants depending on the definition. Many destination countries grant citizenship to foreign-born people who are family members of citizens or who satisfy certain legal and residence requirements. These naturalized citizens continue to be recorded as migrants under the foreign-born definition but not under the foreign citizen definition. Many destination countries (for example, the United States) grant automatic citizenship to people born within their territory regardless of parents' citizenship. Yet others, such as Japan, require at least one parent to be a citizen for children to acquire citizenship, even if they were born within its borders. Because of these differences in citizenship and naturalization laws, the numbers of migrants will be substantially higher in the United States if the foreign-born criterion is

[3] This section highlights many of the nuances in the data, but for fuller treatment of the subject, see Bilsborrow and others (1997).

used. In Japan, on the other hand, the number of migrants comes out higher under the foreign citizenship criteria.

Where data are available for both definitions, priority is given to data by country of birth, for several reasons. First, country of birth is more appropriate in analyzing physical movements and handling the cases of former colonies and dependencies.[4] Second, while nationality can change, place of birth cannot.[5] Third, naturalization rates vary enormously across destination countries. Differences in laws on citizenship criteria (for both migrants and their children born in the destination country) do not affect data based on place of birth. Fourth, when migrants cannot be assigned to a specific origin, they are often recorded under an aggregated umbrella heading. These categories embody ambiguity about a migrant's origin, and since migrants are assigned to aggregated headings more frequently when the citizenship definition is used, the foreign born concept is again favored. Last, for migrants living in disputed territories, such as Kashmir and Western Sahara, an individual's status or official citizenship may be unclear, while country of birth is usually more certain.

How Are Migrants Recorded?

Destination countries employ a wide range of tools to enumerate migrants, including population censuses, population registers and registers of foreigners, border statistics, and worker and residence permits.[6] This article focuses on census and population register records, which are widely available, have the broadest geographic coverage, and include similar questions, thereby yielding more standardized responses. For these reasons, they are the primary sources for most data in the Global Migration

[4] This discussion of definitions highlights the somewhat paradoxical possibility of individuals being classified as migrants without ever having moved across an international border. As mentioned, this is generally possible only in the case of people born in one country but who are citizens only of another country. A similar situation arises with dependencies and former colonies. Residents of Martinique, a French dependency, are automatically granted French citizenship. The statistics for Martinique show all the domestic population as French, possibly leading one to think that Martinique is part of metropolitan France or that most of the population moved to France. In such cases, having data categorized by both foreign born and foreign nationality would enable differentiating between the number of locally born inhabitants of Martinique who are French (referred to as Martiniquais), those born in metropolitan France who moved to Martinique, and people from other countries.

[5] Of course the country of birth may be redefined, as elaborated in the next section.

[6] This article deals exclusively with migrant stocks. Nothing can be gleaned therefore about when a migration took place, save for inferences that can be made by comparing differences in stocks over time. Nor is anything known about the circumstances (such as visa type) under which an individual entered a particular destination country.

Database. Where both censuses and population registers are available, censuses receive priority.

Censuses, generally conducted decennially, are retrospective tools for surveying an entire population (or in some cases, a representative sample) at a single point in time. In addition to their universal coverage, their greatest strength is the inclusion of questions on place of birth and nationality. Censuses also typically aim to enumerate the resident population, whether documented or undocumented (Bilsborrow and others 1997). So although some migrants have a strong incentive to provide false information to enumerators, many undocumented migrants will be captured in these matrices.[7] The size and scope of the census questionnaires vary enormously, both over time and in different destination countries. And there is potential variation in the quality of censuses both across countries and over time. Richer countries have many resources at their disposal to design questionnaires, train interviewers, employ statisticians, and disseminate results. Researchers have little choice but to accept the data at face value. However, where the underlying census is clearly substandard (when there are errors that are obviously not coding errors or not easily corrected), these data are omitted from the analysis.

Popular in many parts of Europe, population registers are continuous reporting systems providing up-to-date demographic and socioeconomic information for everyone surveyed. Typically, registers have evolved over time (from parish records, for example). They were never developed specifically to record international migration information, and they vary considerably across countries. For example, the laws under which individuals are classified as migrants and the conditions under which they are inscribed or deregistered differ greatly (Bilsborrow and others 1997).

The Raw Data

The Global Migration Database is a vast collection of destination country data sources detailing migrant stocks from numerous origin countries and regions (United Nations [2008]). Compiling and maintaining the underlying primary sources require herculean efforts to scour the key census collections of the world and enter the data manually. In total, the database comprises records from some 3,500 separate censuses from more than 230 migrant destination countries and territories, by sex and age. Destination countries make numerous revisions between census waves,[8] and the database incorporates as many of these revised figures as possible.[9]

The starting point is to choose the most relevant source for each destination country from each completed census round.[10] Priority is given to data that are superior

[7] The extent to which illegal migration is captured remains unknown.

[8] Census results are also often released in waves, typically beginning with preliminary estimates and following with incremental releases of more detailed data.

[9] The raw data are available at http://esa.un.org/unmigration.

[10] Bhutan, Colombia, and El Salvador did not conduct censuses during the 2000 round; the relevant censuses for 2005 or 2007 are included instead. Similarly, for seven countries without 1960 censuses, data from the 1950 census round are included. In these

bilaterally and disaggregated by gender.[11] Of the 3,500 sources detailed in the overarching Global Migration Database, 1,107 were suitable for analysis once repeated censuses were removed or combined. Of these, 951 record data disaggregated by gender, as reported in table 1.

{Table 1 about here}

Despite the large number of primary sources, there are still inevitable gaps (table 2). This might be because a particular destination country did not conduct a census in a given decade or disseminate the relevant bilateral or gender-specific information. The majority of the migrants omitted from these censuses are in the Middle East and Africa. The countries of the Middle East are often reticent about releasing data, while many countries in Africa have a long history of conflict. Nonetheless, the 68 countries for which there are complete data account for 68 percent of the world migrant stock in 2000. The 17 countries for which there is only one census account for less than 2 percent of the total stock. The data for earlier decades reflect an identical pattern.

{Table 2 about here}

II. HARMONIZING THE MATRICES

Given the complexities of the underlying data, several major challenges arise in constructing global bilateral migration matrices. The most critical were explained above. In some cases, there is no choice but to recognize that the underlying processes that generated the data are less than ideal and to accept the data at face value. In others, every effort has been made to standardize the data.

Defining the Master Country List

Over the period covered by the 1960–2000 censuses used to construct the global bilateral matrices of migrant stocks (1955–2004), the global political landscape underwent fundamental changes. Many countries, especially in Africa, Oceania, and the Caribbean, gained their independence. Following the end of the cold war, many countries redrew their political boundaries. Some fragmented into smaller nation states, such as the Soviet Union, Czechoslovakia, and Yugoslavia, and others reunified following an extended period of separation, such as Germany and Yemen.[12]

---

cases, each origin countries' migrant stock as a share of the total is calculated in 1950 and these shares are applied to the 1960 total.

[11] There is little standardization in the age brackets that countries use to record migrants' age. This is the main reason why an analysis of migrants' age is omitted from the current study.

[12] Small border changes and territorial disputes are ignored.

A single standard set of countries is specified for the entire timeframe of the database, for both origin and destination locations, so that migration numbers for pairs of countries can be compared over time. Since many new origin and destination countries emerged during the study period, the most current set of countries and regions was chosen.

A region is defined as any geographic entity that conducts its own census and that commonly features as an origin in the others' censuses. For example, Western Sahara is omitted because it does not conduct a census although it is a commonly designated origin region. In all, 226 countries, territories, and regions are included in this list in each of the five migration matrices (see appendix 1). One implication of these inclusion decisions is that migration from Croatia to Germany, for example, is reported in every matrix, even though Croatia did not exist in the early time periods. Researchers interested in migration from Yugoslavia to Germany in 1960 would simply total the individual migration levels from the successor states of Yugoslavia. Performing the analysis according to historical boundaries, though easier, would have masked many recent international movements. Moreover, drawing conclusions about destination countries that no longer exist would offer policymakers less useful information for drawing inferences.

Another complication is the 11 additional destinations with census data that do not map perfectly to the master list. Five of these were aggregated into other countries in the master list: Christmas Islands (to Australia), Cocos Islands (to Australia), Kosovo (to Serbia and Montenegro), South Yemen (to Yemen), and West Germany (to Germany). Six additional countries or territories no longer exist, but they map to two or more of the 226 locations on the master list. These are the Gilbert and Ellice Islands, the former Yugoslavia, Czechoslovakia, Ruanda-Urundi, the Trust Territory of the Pacific Islands, and the Soviet Union. Except for the Soviet Union, the census data for these countries or territories are disaggregated and distributed among the destination countries currently in existence on the basis of more recent migration figures.[13] All of these assignments are made according to the distribution of immigrants of the successor countries in later years.

The Soviet Union is a unique challenge. As mentioned, the enforcement of new borders and the creation of new nation states typically create new migrants overnight. According to the foreign-born definition, people who cross new borders that are created with the break-up of a country are considered migrants, even if they moved before the break-up while the country was still unified. This is particularly problematic in the case of the Soviet Union because 15 new sovereign nations were created overnight, there have historically been large numbers of internal migrants, and migrants have traditionally been recorded using a definition based on ethnicity. Failing to make any adjustment for the Soviet Union, therefore, would result in a large artificial jump in the number of migrants at the time of break-up (see appendix 3).

[13] For example, the 1988 census data for the Trust Territory of the Pacific Islands were disaggregated and distributed among the Marshall Islands, the Federated States of Micronesia, the Commonwealth of the Northern Mariana Islands, and the Republic of Palau. However, in years when a country conducted its own census but was also included in the census of a more aggregated region, the country's own census is prioritized.

Last, specific adjustments are made in the case of Germany and the Republic of Korea. For Germany, bilateral data are available only by nationality. However, these data fail to take adequate account of the large number of ethnic Germans who arrived from other countries between 1944 and 1950 (mainly expellees) and those who arrived after1950 (mainly resettlers). Material from the German 2005 micro-census was therefore used to supplement the data for Germany (see appendix 3). In the case of Korea, data by nationality are readily available for each census round. However, these data fail to account for the large numbers of migrants from the People's Democratic Republic of Korea living in the Republic of Korea. Since the United Nations *Trends in International Migrant Stock* details the total migrant stock in the Republic of Korea by the country of birth definition and because citizenship is rarely granted to people from outside, it is simply assumed that the nationality data were comparable to the foreign-born definition. The nationality total was then subtracted from the UN total and the remaining migrants were assigned to the People's Democratic Republic of Korea.

Recording and Recoding

There is little standardization in the recording and dissemination practices for censuses across destination countries.[14] The level of detail with which destination countries record and disseminate migration data depends on the design of the original questionnaire. Some census questionnaires ask for a specific country of birth and others simply ask for a general geographic region, such as Africa. Even if the original questionnaire asked detailed questions, some countries disseminate data only on how many residents were born abroad or have foreign citizenship. In general, three types of migrant origin are observed in the disseminated census data:

- *Specific geographic regions:* Some of these correspond to exactly one of the 226 countries and territories in the master list. Others pertain to localities that tend to be obscure territories, islands, or regions, such as the Isle of Man or Ceuta.
- *Aggregate geographic regions:* These correspond to two or more countries or territories in the master list. They can be continents (such as Africa), parts of continents (such as South Asia), political alliances (European Union), or other classifications (such as Other Ex-French Africa; Algeria, Tunisia, and Morocco; and Melanesia). The data for these aggregate regions need to be allocated to the 226 countries in the master list. The details of the procedures are discussed below.
- *Miscellaneous categories:* These include refugees, stateless, and born at sea. There are generally no geographic correspondences for these.

Thousands of geographic regions and categories emerged from the more than one thousand individual destination country sources chosen for the analysis. The vast majority of these are repetitions that refer to identical geographic locations using different

[14] The United Nations (1998) has developed recommendations aimed at promoting standardized recording practices across countries. Until such practices are followed uniformly, harmonization will remain a key issue in understanding and comparing migration statistics.

expressions. For example, French Upper Volta and the Republic of Upper Volta were relabeled Burkina Faso. In the end, 292 specific geographic regions (first bullet above) and 236 aggregate geographic regions (second bullet) were identified. The 292 specific regions include the 226 countries and territories in the master list and 66 other single locations that can be assigned to one of the 226 in the master list (see appendix 2).[15]

The 236 aggregate geographic regions pose larger problems. The migrants originating from a given aggregate geographic area need to be allocated to the individual countries that comprise that area. This is one of the greatest difficulties in this project, and resolving it is one of the main contributions of this work. Several propensity measures were developed depending on the quality of the data. They are based either on a destination country's propensity to accept migrants from a particular origin or on origin countries' propensity to send migrants abroad. These propensity shares are then calculated, and the resulting number of migrants are assigned, in order of quality, to specific origin countries in the master list.

Finally, the miscellaneous categories also needed to be dealt with consistently to enable meaningful comparisons between country pairs. There is often a high number of nonresponses to the question about place of birth for foreign-born residents (Bilsborrow and others 1997, p. 60). As a result, some censuses report large number of people with unknown place of birth. All these individuals are assumed to be natives in the analysis since it is unclear as to whether these persons refer to the domestically born or the foreign born. These entries are therefore deleted from the matrices. In other cases, calculations were made to check whether these totals contributed to the foreign born in each census. In most circumstances they did not, and so they were dropped. In cases when these totals did refer to migrants, they were treated as an appropriate aggregate category to be assigned later, as detailed below. Finally, all categories referring to the "stateless"[16] were dropped because despite their importance as a minority group in global migrant patterns, there is no way to meaningfully assign them to an origin.

Disaggregation of Aggregate Categories

The disaggregation of the 236 origin regions identified in the censuses is one of the key steps in creating the bilateral migration matrix. Three propensity equations are used to allocate migrants to one of the 226 countries in the master list. Each measure varies in quality depending on the availability of underlying data. The preferred option is to use migration data from the destination country for the relevant year. If this option is not available, information from the destination country for other years is used. Should

[15] For example, the Vatican is assigned to Italy, Wake Island to the United States, and Labuan to Malaysia.

[16] Some estimates put the number of stateless people (those lacking any citizenship) as high as 11 million, although many of these people will not be captured in censuses. The stateless represent an important category of migrants; for more information, see www.unhcr.org/pages/49c3646c155.html.

that not be possible, subregions[17] are created, and countries with insufficient data are assumed to have a similar propensity to accept migrants as other countries in the subregion. Failing this option, global propensity measures are constructed.[18] More than a single method of allocation is chosen so that the data already in the matrices can be used to maximum effect. All these allocations ignore the gender profile of migrants. This dimension is accounted for at a later stage, once all the aggregate categories have been assigned.

Varying Survey Dates

During the 10-year window of each census round, there are no conventions on when a destination country should conduct its census. Although many destination countries conduct their censuses at the turn of the decade, the actual date is up to each country. Attempting to standardize census dates would require changing the numbers reported in the original census documents.

Most destination countries conduct their census within two years of the middle year of each census round—between 1998 and 2002 for the 2000 census round, for example (table 3). The census numbers thus are not changed, and the matrices report all censuses as comparable in each round. A full list of census dates is in appendix 1. An alternative version of the database that has been mapped to the United Nations (2006, 2009) Trends in International Migrant Stock database is available from the authors. These data are standardized over time in terms of the years to which they refer.

{Table 3 here}

Calculating Missing Gender Splits

Although common in the underlying data, bilateral migration data disaggregated by gender are sparser than aggregate migrant totals (see table 1). An important contribution of the current work is in estimating the gender breakdown of all migrants in destination countries in the global migration matrices. Similar to the allocation from aggregated categories in the Global Migration Database to specific origins in the master list, two measures are used for calculating gender splits; they are described in appendix 5.

Combining Migrant Definitions

Only a single definition of a migrant (foreign born or foreign citizen) can be applied to each destination country in the final matrices. Switching definitions over time

[17] The subregions used for the disaggregations are the 21 UN regions (see http://unstats.un.org/unsd/methods/m49/m49regin.htm, with the countries of Oceania aggregated into a single subregion. They do not match the large World Bank regions used in the analysis in section IV.

[18] While this propensity measure is clearly inappropriate, less than 1 percent of all migrants and observations are assigned on this basis. This method is included so that every migrant in the underlying data is accounted for.

for the same destination country would yield inconsistent data. Priority is given to the foreign-born definition, and these data are always used if at least three censuses using that definition and with detailed bilateral information are available for a particular country. However, only nationality data are available for many destination countries. For countries such as Japan that rarely offer citizenship to foreigners, this does not pose much of a problem since foreign-born and nationality data will be very similar. For other destination countries, including data based on the nationality concept will lead to disparities. When fewer than three foreign born data sources are available and the nationality data are of superior quality, the nationality definition is chosen (see appendix 1). Where fewer than three data points by either definition are available, several assumptions are made to fill the missing data.

Missing Censuses and Census Data

The final hurdle in constructing the global migration matrices is dealing with omitted data. No census round is truly complete since no round has ever included every country in existence at the time. Censuses are expensive because of their universal coverage and labor intensity. For those reasons, many countries have started to conduct censuses only recently (Bhutan began in 2005). Censuses can also be abandoned because of civil unrest or military conflict. They can also be politicized, because they can be used to estimate the size of a particular ethnic group. In other words, data may simply never be released even if they are collected. Nor is there any guarantee that a question on nationality or country of birth will even be included in the census questionnaire. Many countries in Central Asia, as well as Fiji, Sri Lanka, and Tonga, have in some years included questions on ethnicity instead, which is useless for identifying migrants. For all these reasons, inevitable gaps in the data emerge (see table 2).

Three conventions are adopted for constructing missing data. The one that is ultimately used depends on how many data are missing and for which decades these data are missing relative to the decades for which data are available.

MISSING IN-BETWEEN DECADES. Where data are missing for a particular decade but are available for the decade before and after, a linear trend is assumed between the earlier and later bilateral data. In total, 86 country-years of data were interpolated using this method.

MISSING BEGINNING OR END DECADES. Where the data are missing at the beginning or the end of the time period, the destination country is assumed to have the same bilateral migrant composition as in the decade closest to the missing period. The bilateral shares from the closest decade for which data are available are applied to the destination country's total number of migrants for the missing decade. The information comes from one of two sources. In some cases, the census provides the total number of migrants without any bilateral information. If these data are not available, the total from the closest decade is taken and adjusted for growth in migration. The growth rates are taken from Trends in International Migrant Stock, which details total migrant stocks for

all countries in the world at five year intervals (United Nations 2006).[19] The missing end decades are calculated for 116 countries for which data are lacking, most of them for the 1960s and 1970s.[20] Trends in International Migrant Stock database thus can be used to estimate growth rates by estimating missing totals in years for which censuses are not available, and it provides a consistent set of totals over time for countries for that have data of insufficient quality.

An important difference between the matrices presented in this article and the Trends in International Migrant Stock database is the treatment of refugees. While refugees are generally enumerated in developed country censuses, this is not always the case for developing countries. Refugees interned in camps are less likely to be surveyed at the time of census. Making allowances for these refugees, the Trends in International Migrant Stock database adds to the number of migrants refugees reported by the United Nations Refugee Agency and the United Nations Relief and Works Agency for developing countries that are not likely to have included the refuges in their census data. Since the majority of developed countries record refugees alongside other migrants on a bilateral basis, there are normally no remedial measures for removing them. Similarly, for developing countries for which no census data are available, it is impossible to know whether the numbers contained in Trends in International Migrant Stock database include refugees. For the cases that rely on the Trends in International Migrant Stock database, the number of refugees is subtracted from the totals, with the intention of removing refugees in camps from the total, since the focus is on economic migration.[21]

COUNTRIES WITH VERY POOR DATA. For the 59 destination countries for which there are two or fewer census data points, it is impossible to meaningfully interpolate missing census totals or bilateral numbers. In these cases the census totals detailed in the Trends in International Migrant Stock are used. This has the advantage of ensuring consistent totals for the number of migrants in each of the five census periods. The average bilateral shares from the censuses with data are then applied to these totals to derive bilateral data for each census round.

Finally, there are six destination countries for which bilateral data are completely lacking.[22] In these cases, data for all the other countries in the subregion are used to

[19] The 2008 revision includes data only for 1990–2010. To ensure consistent figures over time, the 2005 revision, which covers 1960–2005, was used instead.

[20] Taiwan, China, and Norfolk Island pose an additional problem, since the United Nations does not provide data for these locations, so migrant totals in other years cannot be calculated. For these two areas, therefore, the numbers of migrants are set to zero in the earlier decades for which data are lacking.

[21] In the case of Palestine, for which the UN totals consist entirely of refugees, these totals are not removed. It is possible to calculate migrant totals for Palestine in other decades.

[22] The six countries are China, Eritrea, Maldives, Qatar, Somalia, and Democratic People's Republic of Korea. Of these, , Eritrea, and Somalia have been affected by conflict. China has conducted censuses over the period, but their definition of migration is not compatible with the definitions used throughout the article.

calculate the propensity of every country in the destination subregion to accept migrants from elsewhere in the world. All of the propensities sum to one. These shares are multiplied by the total migrant stock figures provided in the Trends in International Migrant Stock database to calculate the bilateral numbers.

III. RELIABILITY OF THE ESTIMATES

The previous section described the challenges in constructing the matrices and the range of measures used to generate the missing observations. This section highlights the extent to which the estimates are based on the underlying raw data and their reliability.

Categorizing the Methods Used

Nine main methods were used to generate the cells:

1. *Pure raw:* Derived directly from the raw bilateral census data.
2. *Raw scaled:* Based on the underlying raw bilateral data scaled to the UN numbers.
3. *Pure remainder:* Assigned directly from the disaggregation of aggregate categories applying one of the propensity measures.
4. *Remainder scaled:* Based on disaggregations using one of the propensity measures and then scaled to the UN numbers.
5. *R&R* [*Raw and Remainder combined not scaled:* Based primarily on bilateral raw data and to which disaggregations of certain aggregate categories were added.
6. *R&R* [*Raw and Remainder combined*] *scaled:* Similar to R&R not scaled except that the resulting value was scaled to the UN numbers.
7. *Pure interpolation:* Calculated solely by interpolating missing end and middle censuses, but not scaled to the UN data.
8. *Interpolation and scaled:* Both interpolated and scaled, for countries with poor data or for cells calculated by interpolating missing and end decades which then had to be scaled.
9. *Missing:* For countries for which bilateral data were missing for every census round, such as Somalia.

The data used in the first six methods are from the raw census data. The data for the last three methods are missing because of omissions in the underlying data and need to be filled. Therefore, varying percentages of observations in each decade are assigned by the methods described (table 4). In 1960, 59 percent of observations are directly assigned from the raw bilateral data or from one of the disaggregations of the aggregate raw data (the first six categories). By 2000, this proportion rises to 69 percent. However, these observations account for some 84 percent of the total number of international migrants in 1960 (table 5). This proportion rises to 86 percent by 2000 because a small number of corridors (cells) account for a large proportion of global migration stocks. The bulk of the remaining international migrants are assigned on the basis of interpolation.

{Tables 4 and 5 here}

Among the first six categories that are based on raw census data, three categories (raw scaled, R&R not scaled, and R&R scaled) are constructed through the summation of bilateral raw numbers and disaggregations of some aggregate categories in the original censuses. Since these categories together constitute around 45 percent of migrants in each census round, the original bilateral portion of each cell was compared with the final number assigned to them after the various calculations as a check on accuracy. For each decade, therefore, the overall percentage contribution of the raw bilateral data to the total is calculated (table 6).[23] In each census round, at least 92 percent of all those categories are derived from the raw data.

{Table 6 here}

Simulating Missing Data

Finally, to examine the reliability of the estimated missing census data and test the methodologies, several scenarios are assumed. All bilateral observations for a single year for four countries (Australia, United States, Switzerland, and Chile) in different parts of the world are deleted and the missing cells are filled using one of five methods.[24] The first simulation assumes that all bilateral data for 2000 are missing but that the total number of migrants is available. The missing bilateral numbers then have to be filled using the propensity measure (equation 1 in appendix 4) based on the data available in other years. The second and third simulations assume that the total is missing as well, and interpolation is used to fill in all missing data for 1960 and 2000. The fourth and fifth simulations remove all data for all years and then fill the missing years using data for the remaining portion of the subregion (table 7).

{Table 7 here}

The simulations perform well. The four countries are examined one at a time, starting with Australia. The correlation coefficient between the predicted and actual data in each simulation is at least 0.945. Interpolating the data is the most accurate method of predicting the missing data, and simulation 2 for 1960 is more accurate than simulation 3 for 2000. Simulation 1 does not perform as well: the data from other years fail to adequately account for the fairly significant shift in the composition of the Australian immigrant stock after 1990. When simple subregional shares are used (simulations 4 and 5), the correlation coefficients remain high. The actual distribution of immigrants, however, is less accurate, especially in simulation 5. This is because New Zealand, the country in the subregion that has by far the greatest weight for apportioning migrants for

[23] Although only aggregates for each decade are presented here, a full matrix detailing exactly how each cell was generated is available from the authors.

[24] For all countries, data quality is highest for 2000 and lowest for 1960, except for Chile, for which 1980 has the worst quality data.

Australia's missing data, did not experience the same influx of migrants from Asia that Australia did. In other words, Australia represents such a large share of immigration in Oceania that when it is removed, the remaining countries (mostly small island countries that are origins, not destinations) are not particularly accurate predictors of migration to Australia.

The U.S. case is similar. Using interpolation to fill in the missing years proves effective, while the results from simulation 1 are also reasonable. The results from simulations 4 and 5 are less accurate. The problem with using regional shares for calculating missing coefficients for the United States is similar to that for Australia. The poor results are due to the differences in the migrant profiles of the United States and Canada, which provides the weights for filling in the missing U.S. values. This methodology significantly underpredicts the numbers of migrants from U.S. dependencies, since Canada hosts very few of them, and overpredicts the numbers from former British colonies, populations that are more prominent in Canada.

Simulations 4 and 5 perform extremely well for Switzerland: the deviations from the actual data are less than 1 percent. This is due to the fact that several large Western European nations have similar migrant profiles to Switzerland, unlike the case for Australia and New Zealand and the United States and Canada. The data for 1970–2000 prove better for interpolating the missing data for Switzerland for 1960, while the data for earlier years are somewhat less effective at predicting the missing data for 2000.

The results for Chile are also good. Using the data for Chile in other years and the propensity measures yields a margin of error that is under 6 percent (simulation 1). Interpolation proves accurate when data for either 1960 or 200 are removed. With subregional shares, the differences in the log ratios are small, but the correlation coefficients are not as high as in other cases because Chile's immigrant profile is bimodal. Chile has a small number of large immigrant stocks and a large number of very small stocks. Although the predictions for the size of the stocks are reasonable, the relative rankings are not as accurate.

The results indicate that interpolation is the most effective method of allocation, although the allocations based on the propensity measures and on the subregional shares fair reasonably well. This is heartening, since around a quarter of the observations and 14 percent of the world migrant stock is allocated for 2000 using interpolation. Filling a missing country-year of data using propensities is less effective. Even so, the correlations remain high and the resulting data are not sufficiently inaccurate to warrant throwing them away. It is important to remember, however, that simulation 1 represents a worst case. This extreme measure is resorted to only for a few countries for which data are missing. In almost every case, aggregate categories are much narrower in the raw data. Nevertheless, even with this constrained method with extreme assumptions (missing all data for a country in a region with very few comparable countries), the results seem reasonable. And even when the results are skewed, this is generally due to the over- or underpredicting of a handful of key migrant corridors.

Finally, the aggregate figures obtained are compared with those from the Trends in International Migrant Stock database (United Nations 2006, 2009) to highlight key differences. The database provides data by destination only, not for each bilateral

corridor, so only aggregate numbers can be compared. For this comparison, mid-year estimates of the world migrant stock for 1990–2000 are taken from the 2008 edition and estimates for the earlier censuses, 1960–1980, are taken from the 2005 edition (table 8). The analysis subtracts the estimated number of refugees from the total mid-year estimates of the world migrant stock from the Trends in International Migrant Stock database to yield the net number of migrants in each decade. These numbers are then compared with the decadal estimates generated through this project, both the total and the net, after subtracting estimates of migrants within the Soviet Union for 1960–1980 (data for 1990 and 2000 should be directly comparable) and the number of ethnic German migrants added to the German censuses.

{Table 8 here}

The aggregate estimates are remarkably close (the two net totals), differing at most by around 1 million migrants, except in 1990. There are several possible explanations for these differences. First, the census totals from the current work may not match because censuses do not always make allowances for temporary workers. For example, Singapore's official 2000 census records 563,430 foreign-born migrants. The United Nations, however, reports 1,351,806 foreign-born migrants for 2000. Second, there are cases where the current study reports data by nationality, but the corresponding figure in the Trends in International Migrant Stock refers to the foreign born. This situation generally arises when a census does not report the number of foreign-born migrants on a bilateral basis. Examples include Austria and Côte d'Ivoire. Third, differences in the years to which the data refer can generate large disparities. For example, this study uses the 1966 data for Australia, whereas Trends in International Migrant Stock reports data for 1970. Overall, however, the fact that the totals are remarkably close in every decade adds credence to the estimates here.

IV. The Evolution of Global Bilateral Migration

The greatest strengths of the global migration matrices are their bilateral coverage, the number of decades covered, and the disaggregation by gender. These data are too rich for a full analysis of all movements between all pairs of countries. Instead, this section summarizes the major trends in the evolution of bilateral migrant stocks, based primarily on World Bank regions.[25]

Global Trends

The migration matrix for the 1960 census round reflects a realigning world in the postcolonial era. Over the 1960-2000 period, the composition of world migration

[25] Appendix 1 details the World Bank regions: South Asia, East Asia and Pacific, Sub-Saharan Africa, Latin America and the Caribbean, Europe and Central Asia, and Middle East and North Africa. High-income Middle East and North Africa refers to the predominantly oil producing countries in the Persian Gulf (Bahrain, Kuwait, Oman, Qatar, Saudi Arabia, and the United Arab Emirates) and to Israel.

fundamentally changed, driven by world events and increasingly selective immigration policies in developed countries, which led to greatly diversified migrant stocks. Mirroring this pattern, most countries now send migrants to an increasing number of destinations. Migration to developing countries has been driven largely by the partitioning of India[26] and the breakup of the Soviet Union, both events that need be reconciled when interpreting the data. However, while the United States and Western Europe remained throughout the most important destinations, there have been significant migration movements to the other countries of the 'New World' (Australia, New Zealand, and Canada) as well as to the oil-rich Persian Gulf countries (primarily from South and East Asia), reflecting a huge increase in demand for labor following the oil shocks of the 1970s.

Between 1960 and 2000, the total global migrant stock increased from 92 million to 165 million.[27] At the beginning of the period, one fifth of the world's migrant population was born in Europe, and one sixth was attributable to the partition of India and migration within the Soviet Union. Two-thirds of the growth up to 2000 was due to migrant flows to Western Europe and the United States, and the rest was due mostly to increased mobility between the countries of the former Soviet Union, the emergence of the Gulf States as key migrant destinations, greater intra-Africa migration flows, and migration to Australia, New Zealand, and Canada. The number of migrants in South Asia fell over the period, reflecting a falloff after the migrations that followed partition (see figure 2 later in this article). Despite the sustained increase in the global migrant stock over the period, migrants declined as a share of the world population between 1960 and 1990 (from 3.05 percent to 2.63 percent), then rose again slightly to 2.71 percent in 2000.

The importance of migration for destination and origin countries depends on the size of the migrant stock relative to the population. As might be expected, many countries with the highest concentrations of immigrants are small countries with comparatively few people. The countries or territories with a population or more than 1 million people and immigrant ratios over 20 percent in 2000 include the United Arab Emirates (41 percent), Kuwait (38 percent), the Occupied Palestinian Territories (31 percent), Israel (25 percent), and Oman (20 percent). Countries with immigrant ratio less than 1 percent include Indonesia, Madagascar, and Cuba. By destination subregion, migration has become more concentrated in all developed country regions and less concentrated in many developing country regions, especially South and Southeast Asia, South America, and Southern, Eastern, and Central Africa.

Emigration ratios (ratio of emigrants to the sum of the emigrant and domestic populations) were calculated for origin countries. Unsurprisingly, small island states and those experiencing political upheaval or environmental catastrophe have the highest emigration concentrations. In 2000 these included Niue (80 percent), Tokelau (64

[26] It is not possible to differentiate among migrants who moved before, during, or immediately after the partition of India because these migrations occurred before the beginning period of the matrices.

[27] This increase would be starker had it not been for the special treatment of the Soviet Union.

percent), Montserrat (56 percent), Cook Islands (53 percent), and Palau (47 percent). Countries or territories with more than 1 million residents and the highest emigration concentrations include Jamaica (26 percent), the Occupied Palestinian Territories (24 percent), Albania (23 percent), Bosnia and Herzegovina (23 percent), Republic of Ireland (23 percent), and Armenia (22 percent). Those at the other end of the spectrum include Mongolia (2 percent), Madagascar (4 percent), Ethiopia (4 percent), and Brazil (5 percent). By subregion of origin, emigrant concentrations have remained far more stable over the period than immigrant ratios across most of the world. Notable changes have occurred, however, in emigration ratios in the Pacific and the Caribbean and Central America (higher) and South Asia (lower).

Global Migration between the 'North' and the 'South'

Dividing the world into two regions, the North (developed countries) and the South (developing countries),[28] highlights some of important patterns underpinning international migration over the second half of the twentieth century. The number of migrants from the North remained fairly stable, while the number from the South increased (figures 1 and 2). Much of the growth in the number of migrants is driven by migrations from the South to the North, which rose from 14 million to 60 million between 1960 and 2000.

{Figure 1 here}

{Figure 2 here}

Numerically, South-South migration dominates global trends, although this migration is declining as a proportion of total world migration. In 1960, South-South migration accounted for 61 percent of the total migrant stock; by 2000, it had fallen to 48 percent. When the migrant-creating effects of South Asia and the Soviet Union are factored in, however, South-South migration remains stable over the period, at about a quarter of the total (see figure 2). As a proportion of total migrant stock, only South-North migration rose between 1960 and 2000. Increasingly liberal immigration policies in developed countries have been paralleled by large movements from developing countries. The data show that the proportion of world migration attributable to South-North migration rose from 16 percent to 37 percent. This dramatic increase is unquestionably one of the defining trends of the period, surpassing migration between developed countries from 1970 to 1980, both in numbers and as a proportion of the total migrant stock.

[28] The developed countries are Australia, Canada, Japan, New Zealand, the United States, and the EU-15 and the European Free Trade Association, which have all been relatively affluent over the entire period of interest. The EU-15, rather than some other European Union grouping, is included because the latest year to which the data refer is 2004. All other countries are classified as developing.

Global Migration to Developed Countries

The growth in the South-North migration has been driven largely by movements to the United States and Western Europe. Between 1960 and 2000, migrant stocks grew by 24.3 million in the United States and 22 million in Western Europe, accounting for some 42 percent of the world total in 2000. However, there are notable differences in the migrant compositions of these two regions. Whereas the U.S. immigrant profile has changed dramatically, Europe's has remained more stable, reflecting in part its continuing ties with former colonies.

Immigration to the United States in 1960 was dominated by Europeans, who accounted for around 60 percent of the total and 6 of the top-10 migrant corridors. Of the 10.4 million migrants in the United States at that time, 1.26 million were born in Italy, 990,000 in Germany, 835,000 in Great Britain, 750,000 in Poland, 360,000 in Ukraine, 340,000 in Ireland, and 305,000 in Austria. By 2000, the share of these origin countries declined, to around 15 percent. Balancing this trend, the number of migrants from Latin America and the Caribbean and East Asia and Pacific rose sharply. In 2000, 52 percent of the immigrant stock in the United States were born in Latin America and the Caribbean and 17 percent in East Asia and Pacific.

The United States is an important destination for migrants from all regions except Southern and Central Africa. In 2000, the United States received the largest number of migrants[29] from 60 countries, including Germany, Vietnam, Cuba, and the Republic of Korea. Moreover, 13 of the 50 largest migration corridors in the world and 6 of the 10 largest South-North corridors in 2000 were to the United States. The two largest corridors to the United States were from Mexico and the Philippines, the largest and 12th largest developing to developed country migration corridors in the world. They accounted for 10.8 million migrants, equivalent to 31 percent of the migrant stock in the United States, or nearly 7 percent of the world migrant stock.

Western Europe has been instrumental in many of the largest migrations in history, as both a major sending and receiving region. Between 1960 and 2000, many Western European countries transformed from net migration senders to net migration receivers. Today, Western Europe remains a key destination region for migrants from every other part of the world except the high-income Middle East and North Africa region. Increasingly over the period, Western Europeans began migrating to other countries in the region. In 2000, two-fifths of Western European migrants lived elsewhere in Western Europe, driven largely by the expansion and economic and political integration of the European Union. This is a significant increase from 1960, when far greater numbers of Europeans chose to migrate to the United States and to Latin America and the Caribbean. Despite these increases, however, intra-Western European migrants are increasingly becoming a minority proportion of the migrant stock, especially after 1970 as migration from developing countries increased. Migrants from Turkey and

[29] Migration corridors are discussed to highlight the most important global migrant stocks; at no point does the discussion relate to migration flows. The focus is on stock data, and the term "migration corridor" simply refers to the bilateral migrant stock for a particular pair of countries.

Poland in Germany constitute the two largest diasporas in Western Europe and the second and third largest developing to developed countries migration corridor globally. Elsewhere in Europe, the most significant migrant corridor from developing countries is from Algeria to France. In all decades except 2000, this corridor is among the top four most important developing to developed country migrations in the world. Other notable corridors from the South to Western Europe include South Asia to Great Britain, the former Yugoslavia to Germany, and North Africa (countries in addition to Algeria) to France.

Modern day Australia, New Zealand, and Canada were all founded through immigration; in 1960, 71 percent of migrants to Australia, New Zealand, and Canada were born in Western Europe—39 percent of them in the United Kingdom. By 2000, however, that share had fallen to 36 percent of the total, as migrants from the East Asia and Pacific region (particularly China and Vietnam) gained prominence; they now account for more than a fifth of migrants.

Germans in the United States and British in Australia are the two largest migration corridors between developed countries. Facing a chronic skills shortage, Australia implemented the Ten Pound Pom scheme in the postwar period as part of its Populate or Perish policy. Opening the country to all British citizens, including those from Cyprus and Malta, the Australian government managed to persuade over one million people to migrate before 1973[30] for the price of just 10 British pounds. Given the cultural similarities between Australia and the United Kingdom and the relaxed reciprocal visa restrictions, bilateral migration flows remain strong to this day. Japan has historically been more reticent than other OECD members to admit migrants. Immigration to Japan is mainly from Korea and elsewhere in East Asia, although from 1960 onwards, Japan did admit larger proportions of migrants from both Southeast Asia and South America, specifically Brazil, the *Nikkei burajiru-jin*.

Global Migration to Developing Countries

Statistically, the most important events affecting migrant movements to the South over the study period are the partition of India and the disintegration of the Soviet Union. There have been other important changes as well since 1960, particularly the large shift in global migration toward the Persian Gulf countries.

In 2000, 15 percent of the migrant stock in developing countries (including both India partition and intra-Soviet Union migrants) was in the high-income Middle East and North Africa region, up from under 3 percent in 1960. These migrants reflect movements predominantly from South and Southeast Asia (45 percent in 2000) and the low-income Middle East and North Africa region (33 percent) to the Gulf and from the countries of the former Soviet Union to Israel.[31] Of total migration to developing countries, the low-income Middle East and North Africa and the Latin America and Caribbean regions continue to attract steady shares. Compared with 1960, however, both regions attract

[30] From 1973 onward, the price of assisted migrant's passage rose.

[31] In 1960, over half of all migrants in Israel were born in the Eastern Europe and Central Asia.

proportionally far fewer Western Europeans and more migrants from other developing countries. Although the number of migrants across Africa increased by some 4 million over the period, Sub-Saharan Africa accounted for only 14 percent of total migrants in developing countries in 2000, down from 11 percent in 1960. The numbers of migrants in Southeast Asia, Europe other than European Free Trade Association and the EU 15, and Eastern Africa fell over the period, reflecting a sharp drop in migrants from East Asia in Southeast Asia, fewer migrants from the former Soviet Union in Eastern Europe, and fewer migrants from South Asia and East Africa to other developing countries in the subregions.

Intra-Soviet Union and intra-South Asia migration constituted 42 percent of South-South migration globally in 2000 (figure 3). The largest migrant corridors were between countries of the former Soviet Union, between Russia and Ukraine (in both directions), and between Kazakhstan and Russia. Migrant corridors between Bangladesh, India, and Pakistan are very large in both directions, with Bangladeshi migrants in India the largest migrant population in South Asia. In the Persian Gulf, the largest migrant groups are Indian and the Egyptian migrants in Saudi Arabia, Indian migrants in the United Arab Emirates, and Pakistani migrants in Saudi Arabia.

{Figure 3 here}

Migration from the North to the South, although still large, is declining (see figure 2). In 1960, developed country migrants constituted the majority of migrants to the Pacific Islands, Central and South America, and Central Africa; today, that is no longer the case. Migrants from developed to developing countries have declined in both absolute and relative importance. Today, the most important developed to developing country movements are from Western Europe to South America and to other European countries and from the United States to Central America and the Caribbean. Migrants from the United States to Mexico constitute the largest developed to developing country migration corridor in the world today, at more than 340,000 people. Before 2000, migration between Italy and Argentina was the largest developed to developing country migration corridor in every decade. Other notable developed to developing country corridors are from Spain to Argentina and from Great Britain to South Africa.

Gender Assessment of International Migrant Stocks

In 1960, men made up a larger share of all regional immigrant stocks except in the United States and Eastern Europe and Central Asia (figure 4). Between 1960 and 2000, the gender composition of immigrant stocks changed considerably. The United States, Eastern Europe and Central Asia, and South Asia all experienced slight declines in the share of women in total migrants. The largest percentage increases over the period in the share of women in the total migrant stocks were Latin America and the Caribbean (14.8 percent); Japan (14.3 percent); East Asia and Pacific (13.3 percent); Sub-Saharan Africa (11.2 percent); Australia, New Zealand, and Canada (8.3 percent); and Western Europe (4.9 percent). The proportion of women in the migrant stock fell sharply in both the high-income Middle East and North Africa region (23.8 percent) and the low-income Middle East and North Africa region (9.1 percent drop) .

{Figure 4 here}

In absolute terms, however, the number of female migrants in all regions but South Asia rose. Despite the high-income Middle East and North Africa region hosting fewer women than men, the region experienced the largest rise in the number of female migrants (up 3.5 million or 540 percent) over the period. Other regions that experienced large increases in the number of female migrants include the United States (up 12.1 million or 228 percent); Western Europe (11.2 million, 190 percent); and Australia, New Zealand, and Canada (3 million, 130 percent). The biggest absolute decline in the numbers of female migrants between 1960 and 2000 was in South Asia (down 3 million or 40 percent). In 2000, the countries with the highest proportion of female migrants were Nepal (70 percent), Mauritius (63 percent), and Moldova (60 percent).

In terms of emigrant stocks in 1960, only two regions sent higher numbers of women abroad relative to men, Australia, New Zealand, and Canada and Eastern Europe and Central Asia (figure 5). They did so again in 2000, along with Western Europe, East Asia and Pacific, and Japan. In percentage terms, the ratio of female to male emigrants declined slightly in the United States; Australia, New Zealand, and Canada; and Eastern Europe and Central Asia and more substantially in South Asia (9.6 percent) and in both Middle East and North Africa regions (high income, 6.2 percent; low-income, 7.8 percent). The four regions that experienced the greatest increases also experienced the largest increase in women as a share of their total immigrant stocks: East Asia and Pacific (17.9 percent), Japan (15.5 percent), Sub-Saharan Africa (15.4 percent), and Latin America and the Caribbean (6.9 percent). In absolute terms, all regions of the world sent more women abroad in 2000 than in 1960. The largest proportional increase was from Latin America and the Caribbean (up 10.9 million or 630 percent), followed by the high-income Middle East and North Africa region (500,000, 290 percent), the low-income Middle East and North Africa region (3.3 million, 250 percent), Japan (330,000, 210 percent), East Asia and the Pacific (6.3 million, 180 percent), and Sub-Saharan Africa (4.4 million, 180 percent). In 2000, the countries with the highest proportion of women in their emigration stocks were Ukraine (61 percent), the Philippines (60 percent), and Singapore (60 percent).

V. CONCLUSION

This article draws on the largest collection of censuses and population registers providing information on international bilateral migration and constructs consistent square matrices for the last five completed census rounds (1960 to 2000). Problems in the underlying data that confound meaningful comparisons include differences in recording and recoding practices among destination countries and missing and omitted data.

The main contribution of this article is in recognizing and overcoming these obstacles by making a series of simplifying assumptions. Tradeoffs between pragmatism and accuracy are inevitable, and one of the largest hurdles is establishing a set of rules for achieving a fixed set of countries. Researchers face daunting challenges when working with migration data, and any attempt to resolve them will inevitably fall short of the

ideal, especially when compared to international statistics on trade and financial flows. Nevertheless, given the paucity of comparable data on international migration, especially outside of the OECD, the completed database represents an important step in an ongoing effort to understand trends in international migration. The matrices provide a reasonably accurate portrait of global migration over the second half of the twentieth century and should provide a useful starting point for researchers and policymakers working on a broad range of issues.

APPENDIX 1. LIST OF SOURCES

{Table A1 here}

APPENDIX 2. LIST OF AGGREGATIONS

{Table A2 here}

APPENDIX 3. ADJUSTMENTS TO THE DATA

This appendix describes the adjustments made to the data for the former Soviet Union and Germany.

Former Soviet Union

Censuses for the Soviet Union for 1959, 1970, 1979, and 1989 were collected to address the data issues created by the dissolution of the Soviet Union. These censuses all use ethnicity to identify migrants. Crucially, for 1989, comparable country of birth data exist for all 15 republics. The censuses based on ethnicity document intra-Soviet migrants (Uzbeks in Turkmenistan, for example) and external nationalities (such as Afghans). In addition, there are miscellaneous Soviet nationalities (such as the Chuvash, Tatars, and Uyghurs), many of whose homelands span several Soviet republics/countries and who should therefore not be counted as international migrants since they were born on one side of the border or the other as opposed to moving across it.

First, people of these miscellaneous nationalities were broadly aggregated to one or more of the 15 former Soviet republics on the basis of country by country research and a close inspection of the numbers over time. Similarly, external nationalities were assigned, with particular attention to determining whether these people were actually migrants. For example, people recorded as Germans will likely be ethnic Germans who migrated long before the census period examined in this study. Those recorded as Poles, however, are more likely to have been forcibly deported. Once the aggregations were completed, the ratios of foreign-born migrants to migrants defined by ethnicity in 1989 were calculated for people who were both born in one of the 15 former Soviet republics and resided there. These ratios were then applied to these republics/countries in every census period before adding the "external" migrants. These corrections captured a large proportion of the most important migrants to and between the Soviet republics. This process adds many millions of migrants to the totals in the early decades and avoids the problem of a very large artificial jump in international migration between 1980 and 1990, after the dissolution of the Soviet Union.

Germany

The 2005 German micro-census includes data on emigrants of German origin from Eastern Europe who arrived between 1944 and 1950 (referred to as expellees, *Vertriebene*) or between 19502005 (referred to as resettlers, *Aussiedler*). These data are recorded by year of birth and year of migration; country of birth is not recorded. As of 1950, there were 11.96 million expellees and 4.48 million resettlers residing in Germany. According to the data provided by the Max Planck Institute for the Study of Religious and Ethnic Diversity, 3.61 million were still in Germany as of 2005. Mortality data from the United Nations Population Division (United Nations, 2010) on Germany for each decade and age group were used to calculate the number of migrants who would have been residing in Germany at the beginning of each decade from 1960 to 2000, taking into account migrants' age and year of entry. After calculating the total number residing in Germany in each decade, shares were estimated by country of origin using the nationality

shares from the 1950 data on expellees and post-1950 data on resettlers. The numbers of expellees and resettlers were then added to the existing totals.

APPENDIX 4. PROPENSITY MEASURES

This appendix presents the propensity measures used to disaggregate the 236 aggregate origin regions/countries identified in the censuses. Let $M_{o,d,t}$ denote the number of migrants from origin country *o* in destination country *d* in year *t*. These are the entries in the bilateral matrices that need to be completed. Now, instead of $M_{o,d,t}$, suppose a census in country *d* gives the number of migrants originating from region *R* (which includes country *o*), denoted as $M_{R,d,t}$. The problem is to find an allocation rule ($\sigma_{o,d,t}$) for estimating the bilateral stock from this aggregate amount. The allocation rule can be written as $M_{o,d,t} = \sigma_{o,d,t} M_{R,d,t}$.

One type of aggregation problem occurs in the case of migrants from Czechoslovakia, the Soviet Union, and Yugoslavia and their successor states. For example, in many cases, migrants are recorded from Czech Republic, Slovakia, and Czechoslovakia in the same year. Belgium's 2001 reports 308 migrants from Czechoslovakia, 554 from the Czech Republic, and 412 from Slovakia. Presumably, migrants who left before the partition reported Czechoslovakia as their origin country, whereas most postpartition migrants reported the successor countries. In such cases, it is assumed that the distribution of migrants from these two countries was the same before and after the break-up of Czechoslovakia. Of the 308 migrants recorded as originating from Czechoslovakia, 177 migrants (308*[554/966]) were assigned to the Czech Republic and 131 (308*[412/966]) to Slovakia.

In other cases of aggregated migrant stock data, migrant data from other decades were used as the basis for disaggregation. Migrants were allocated according to a relative propensity, which is averaged over time. This can be formally written as:

$$\sigma_{o,d,t} = \left(\frac{1}{n}\right)\sum_{k\in K}\sigma_{o,d,k} = \left(\frac{1}{n}\right)\sum_{k\in K}\left(M_{o,d,k} / M_{R,d,k}\right) \qquad (1)$$

where *K* denotes the set of census years other than *t* for which bilateral data exist, and *n* is the number of such observations in set *K*. This propensity is simply the likelihood that a particular destination country will accept migrants from a specific origin country, relative to all the other countries comprising that aggregate origin region. For example, Australia records 29,311 migrants from the Soviet Union in 1966. This total needs to be disaggregated among the 15 successor countries in the master list. While the data for Australia cover census material for each of the five census rounds, only the 2001 census provides details for all 15 successor countries. According to the first method for allocating aggregate categories, the 2001 census is used to calculate the contribution of each of these countries towards the total. Those shares are then used to allocate the

29,311 migrants from the Soviet Union in 1966 among the constituent republics to yield the bilateral numbers for Australia (table A3).

{Table A3 here}

In this simple example, only the data for 2001 are available. Where data are available for more than one census, the shares across all decades are averaged before estimating the bilateral numbers.

In the absence of such data (disaggregated data for the *same* destination country in *other* census years), the world is disaggregated into destination subregions. Origin countries in the same subregion are then assumed to have a similar propensity over time to send migrants to a particular destination country in a subregion for which data are lacking as they do to other countries in that subregion. For example, assume that the census data for Morocco in a particular year include the origin category All West Africa but no individual data on migrants from Ghana and that there are no bilateral data on Ghanaian migrants in other Moroccan censuses. In this instance, migrants from Ghana are assumed to have a similar propensity to migrate to Morocco as they have to other countries in North Africa. Data from other countries in North Africa (Algeria, Egypt, Libya, and Tunisia) are then used to calculate the propensity of Ghanaians—relative to migrants from other West African countries—to migrate to each country in North Africa. These propensity shares, which sum to one, can be applied to the All West Africa aggregate category from the Moroccan census to disaggregate it into the constituent West African countries. Equation 2 expresses this propensity measure:

$$\sigma_{o,d,t} = \left(\frac{1}{nf}\right)\sum_{k\in K}\sum_{g\in G}\left(M_{o,g,k} / M_{R,g,k}\right) \qquad (2)$$

In equation (2), *G* denotes the set of comparable destination countries (Algeria, Egypt, Libya, and Tunisia in the example above); *R* is the set of origin countries (All West Africa); *n* is the number of census years for which data exist; and *f* is the number of countries in region *G*. In short, this is the relative propensity of origin country *o* to send migrants to subregion *G* relative to other countries in its own region (*R*). Where appropriate data for the subregion cannot be found, the set of all countries in the world is used.

Appendix 5. Calculating Gender Splits

When gender splits are missing, the preferred option is to divide the world into subregions. Then it is assumed that the gender ratio of an origin country's emigrant stock in a specific decade is the same for each destination country in that subregion. The missing gender ratio in an origin country's emigrant stock can then be calculated using

data disaggregated by gender from all destinations in the same subregion as the destination country for which data are lacking. Using the same notation as in the previous section, assume that $M_{o,d,t}$ is the aggregate migrant stock from origin country *o* to destination country *d* in year *t* and that $W_{o,d,t}$ is the female migrant stock for the same origin-destination pair in the same year *t*. The ratio of female migrants to male migrants is denoted as $\gamma_{o,d,t}$, which is given by $\gamma_{o,d,t} = W_{o,d,t} / M_{o,d,t}$.

For example, imagine that in a given decade, the gender splits of emigrants from Uruguay in Scandinavian countries are known, except for Sweden. In this situation, it is assumed that the ratio of female migrants to male migrants from Uruguay to Sweden is the ratio of female migrants to male for all of the other Scandinavian countries (Denmark, Finland, Norway) in that decade. Formally, this can be stated as

$$\gamma_{o,d,t} = W_{o,G,t} / M_{o,G,t} \tag{3}$$

where *G* is the destination region (the Scandinavian countries except Sweden), *o* is the origin country (Uruguay), and *d* is the destination country (Sweden). Once this proportion $\gamma_{o,d,t}$ is calculated, it can be multiplied by the total number of migrants $M_{o,d,t}$ to Sweden to calculate the number of female migrants. There is considerable variation in the balance between male and female migration from Uruguay to Scandinavian countries other than Sweden (Denmark, Finland, Norway) during the 1990 census round (table A4). On average, however, 47 percent of Uruguayan migrants are men and 53 percent are women. In the 1990 census, Sweden records 2,640 migrants as originating from Uruguay. Then 1,390 (0.53*2,640) of these migrants are women and 1,250 (0.47*2,640) are men.

{Table A4 here}

These calculations based on concurrent shares can be calculated only if data disaggregated by gender exist for all other countries in the destination subregion. If not, the world is divided into destination subregions, and gender splits are calculated based on regional shares over time. Continuing from the previous example, assume the data for Denmark, Finland, and Norway are unavailable in 1990, so that the gender split for Uruguayan migrants in Sweden cannot be calculated based on Scandinavian data for 1990. In this case, the data for Scandinavia across all other decades are used to calculate the average ratios of female migrants to total migrants over time. This can be written formally as:

$$\gamma_{o,d,t} = \left(\frac{1}{n}\right)\sum_{k \in K} \gamma_{o,G,k} = \left(\frac{1}{n}\right)\sum_{k \in K} \left(W_{o,G,k} / M_{o,G,k}\right) \tag{4}$$

The expression in brackets [$W_{o,G,k}/M_{o,G,k}$?] is the ratio of female migrants to male migrants from origin *o* to all destination countries in the destination subregion *G*, across all decades *k*, for which data exist. Of course, complete data are not available for the current decade *t* since, were that the case, equation (4) would be preferred. Again, once calculated, this share is multiplied by the total number of migrants to determine the number of female migrants.

REFERENCES

Beine M., Docquier, F., and H. Rapoport. 2007. "Measuring International Skilled Migration: A New Database Controlling for Age of Entry." *The World Bank Economic Review* 21 (2): 249–54.

Bhargava, A., and F. Docquier. 2008. "HIV Pandemic, Medical Brain Drain, and Economic Development in Sub-Saharan Africa." *The World Bank Economic Review* 22 (2): 345–66.

Bilsborrow, R.E., G. Hugo, A.S. Oberai, and H. Zlotnik. 1997. *International Migration Statistics, Guidelines for Improving Data Collection Systems.* Geneva: International Labour Office.

Docquier, F., and A. Marfouk. 2006. "International Migration by Educational Attainment (1990-2000): Release 1.1." In *International Migration, Remittances and Development,* ed. C. Ozden and M. Schiff, [151-201]. New York: Palgrave Macmillan.

Docquier, F., B.L. Lowell, and A. Marfouk. 2009. "A Gendered Assessment of Highly Skilled Emigration." *Population and Development Review* 35 (2): 297–321.

Harrison, A., T. Britton, and A. Swanson. (2003. *Working Abroad: the Benefits from Nationals Working in Other Economies.* Paris: Organisation for Economic Cooperation and Development.

Integrated Public Use Microdata Series. [2008]. International: Version 6.0 [Machine-readable database]. Minnesota Population Center, University of Minnesota, Minneapolis, MN.

Mayda A. M. 2007. "International Migration: A Panel Data Analysis Of The Determinants Of Bilateral Flows." CReAM Discussion Paper Series 0707, Centre for Research and Analysis of Migration, Department of Economics, University College London.

OECD (Organisation for Economic Co-operation and Development). 2002. Trends in International Migration database. SOPEMI 2002 Edition. Paris: OECD.

———. 2008. *A Profile of Immigrant Populations in the 21st Century: Data from OECD Countries*. Paris: OECD.

Parsons, C.R., R. Skeldon, T.L. Walmsley, and L.A. Winters. 2007. "Quantifying the International Bilateral Movements of Migrants." In *International Migration, Economic Development and Policy,* ed. C. Ozden and M. Schiff. New York: Palgrave Macmillan.

Ratha, D., and W. Shaw. 2007. “South-South Migration and Remittances.” World Bank Working Paper 102, World Bank, Washington, DC.

United Nations Statistics Division. 1998. Recommendations on Statistics of International Migration Revision 1. New York: United Nations.

United Nations, Department of Economic and Social Affairs, Population Division. [2008]. United Nations Global Migration Database. New York: United Nations. http://esa.un.org/unmigration

———. 2006. Trends in Total Migrant Stock 1960–2000, 2005 Revision. Database. POP/DB/MIG/Rev.2005/Doc. New York: United Nations.

———. 2009. Trends in International Migrant Stock: The 2008 Revision. Database. POP/DB/MIG/Stock/Rev.2008. New York: United Nations. http://www.un.org/esa/population/.

———. 2010. “World Population Prospects: The 2009 Revision, Highlights”, Working Paper No. ESA/P/WP.210, New York, New York

[Fig 1.: Change North to developed countries and South to developing countries in legend]

FIGURE 1: Changes in the Number of Migrants in Developed to Developing Country Migration Corridors, 1960–2000

Source: Authors' calculations based on data described in text.

[Fig 2: Change North to developed countries and South to developing countries in legend; label y axis "Millions"]

FIGURE 2: Changes in the Share of Migrants by Migration Corridors, 1960-2000 (percentage contribution)

Source: Authors' calculations based on data described in text.

FIGURE 3: Inter- and Intra- regional Migration between Developing Countries, 2000

Source: Authors' calculations based on data described in text.

[Fig 4: Write out all region names on y axis; change X axis label to "Percent"]

FIGURE 4: The Percentage of Women in Immigrant Stock by Region, 1960 and 2000

Source: Authors' calculations based on data described in text.

[Fig 5: Write out all region names on y axis; change X axis label to "Percent"]

FIGURE 5: The Percentage of Women in Emigrant Stock by Region, 1960 and 2000

Source: Authors' calculations based on data described in text.

TABLE 1. Total Number of Database Sources

Census round	Birthplace sources	Nationality sources	Total national sources	Birthplace by gender	Nationality by gender
1960	124	67	149	103	64
1970	112	52	133	92	49
1980	145	86	164	117	80
1990	151	114	175	129	99
2000	134	122	161	118	100
Total	666	441	782	589	392

Source: Authors' calculations based on data described in text.

TABLE 2. Number of Missing Census Rounds

Number of missing census rounds	Number of destination countries	Share of world migration in 2000 (%)
0	68	68
1	55	12
2	41	10
3	39	8
4	17	2
5	6	0
Total	**226**	**100**

Source: Authors' calculations based on data described in text.

TABLE 3. Percentage of Censuses Conducted during the Middle of each Census Round

Census round	Censuses by birthplace	Censuses by nationality
1960	78	66
1970	71	71
1980	78	59
1990	80	58
2000	84	57

Source: Authors' calculations based on data described in text.

TABLE 4. Percentage Distribution of Observations by Allocation Method

Census round	Pure raw	Raw scaled	Pure remainder	Remainder scaled	R&R not scaled	R&R scaled	Pure interpolation	interpolation and scaled	Missing
1960	12.31	0.13	40.65	3.34	2.12	0.02	24.86	12.58	3.98
1970	12.07	0.02	34.88	2.03	2.54	0.17	33.64	10.68	3.98
1980	12.00	0.15	45.76	5.90	4.28	0.15	18.88	8.91	3.98
1990	13.82	0.27	47.13	5.88	6.50	0.06	15.23	7.12	3.98
2000	12.85	1.02	39.97	4.29	10.41	0.86	10.66	15.97	3.98

Source: Authors' calculations based on data described in text.

R&R: Raw and Remainder combined.

TABLE 5. Percentage Distribution of Migrants by Allocation Method

Census round	Pure raw	Raw scaled	Pure remainder	Remainder scaled	R&R not scaled	R&R scaled	Pure interpolation	Interpolation and scaled	Missing	Total (millions)
1960	28.61	8.73	7.10	0.51	39.23	0.20	4.72	10.51	0.40	92.3
1970	42.94	0.00	4.01	1.97	34.25	0.25	12.00	4.11	0.46	102.4
1980	30.12	0.23	4.42	0.07	48.77	0.14	11.31	4.64	0.30	118.6
1990	36.61	0.55	4.79	0.35	43.07	0.15	10.05	3.84	0.58	139.4
2000	35.62	1.18	7.56	0.33	40.22	1.07	7.08	6.23	0.72	165.3

Source: Authors' calculations based on data described in text.
R&R: Raw and Remainder combined.

TABLE 6. Contribution of Raw Bilateral Data to the Total

Census round	Accounted for by "raw" data (%)
1960	95.9
1970	92.5
1980	92.5
1990	92.1
2000	93.6

Source: Authors' calculations based on data described in text.

TABLE 7. Five Simulations Testing the Reliability of Generated Cells with Missing Data

	1 Propensity (2000 removed)		2 Interpolation (1960 removed)		3 Interpolation (2000 removed)		4 Missing (1960 removed)		5 Missing (2000 removed)	
Country	Correlation coefficient	Log ratio	Correlation coefficient	Log ratio	Correlation coefficient.	Log ratio	Correlation coefficient	Log ratio	Correlation coefficient	Log ratio
Australia	0.945	–0.575	0.998	–0.067	0.990	–0.126	0.954	–0.242	0.946	–1.302
United States	0.893	–0.243	0.961	0.169	0.972	–0.262	0.596	0.284	0.250	0.124
Switzerland	0.771	0.393	0.971	0.041	0.900	–0.312	0.899	–0.006	0.818	0.010
Chile	0.688	0.059	0.997	0.183	0.897	0.038	0.498	–0.080	0.376	0.009

Note: A cutoff of 250 migrants is implemented for calculating the log ratios since they can be highly skewed by the predictions of very small corridors.
Source: Authors' calculations based on data described in text.

TABLE 8. Comparison of Aggregate Numbers with the United Nations Trends in International Migrant Stock Database

	Unite Nations database			Current study			
Census round	Total	Refugees	Net total	Total	Within the Soviet Union	Germans	Net total
1960	75.5	2.2	**73.3**	92.3	15.8	3.7	**72.7**
1970	81.3	3.9	**77.4**	102.4	21.0	3.8	**77.6**
1980	99.3	9.1	**90.2**	118.6	23.6	3.8	**91.3**
1990	155.5	18.5	**137.0**	139.4	-	4.7	**134.7**
2000	178.5	15.6	**162.9**	165.3	-	3.8	**161.5**

Source: Authors' calculations based on data described in text and United Nations (2006, 2009)..

TABLE A1. List of Database Sources by Census Round

Country or territory	Definition[a]	1960	1970	1980	1990	2000
Australia and New Zealand						
Australia	FB	1961	1966	1981	1991	2001
New Zealand	FB	1961	1971	1981	1986	2001
Japan						
Japan	NT	1960	1970	1980	1990	2000
Canada						
Canada	FB	1961		1981	1986	2001
United States						
United States	FB	1960	1970	1980	1990	2000
Western Europe						
Andorra	NT		1969	1984	1994	2004
Austria	NT	1961	1971	1981	1991	2001
Belgium	NT	1961	1970	1981	1991	2001
Cyprus	FB	1960			1992	2001
Denmark	FB	1960	1965	1981	1991	2001
Faeroe Islands	NT				1994	2004
Finland	FB	1960	1970	1980	1990	2000
France	FB	1962	1968	1982	1990	1999
Germany	NT(FB)	1960	1970	1980*	1990*	2000
Gibraltar	FB	1961	1970	1981	1991	2001
Greece	NT	1961	1971	1981	1991	2001
Iceland	FB	1960	1965	1980	1990	2000
Ireland	FB	1961	1970	1981	1986	2002
Italy	FB	1961	1971	1981	1991	2001

Liechtenstein	NT	1960	1970	1980	1990	1998
Luxembourg	FB	1960	1970	1981	1991	2001
Malta	NT	1957	1967			1995
Monaco	FB	1961	1968	1982	1990	2000
Netherlands	FB	1960			1992	2002
Norway	FB	1960	1970	1980	1990	2000
Portugal	FB	1960		1981	1991	2001
San Marino	NT		1972	1980		
Spain	FB	1960		1981	1991	2001
Sweden	FB	1960	1970	1980	1990	2000
Switzerland	NT	1960	1970	1980	1990	2000
United Kingdom	FB	1961	1971	1981	1991	2001
Eastern Europe and Central Asia						
Albania	NT				1989	
Armenia	ETH(FB)	1959	1970	1979	1989	2001
Azerbaijan	ETH(FB)	1959	1970	1979	1989	
Belarus	ETH(FB)	1959	1970	1979	1989	1999
Bosnia & Herzegovina	FB			1981*		
Bulgaria	FB					2001
Croatia	FB			1981*	1991	2001
Czech Republic	FB				1991*	2001
Estonia	ETH(FB)	1959	1970	1979	1989	2000
Georgia	ETH(FB)	1959	1970	1979	1989	
Hungary	NT	1960				2003
Kazakhstan	ETH(FB)	1959	1970	1979	1989	
Kyrgyzstan	ETH(FB)	1959	1970	1979	1989	1999
Latvia	ETH(FB)	1959	1970	1979	1989	2000

Lithuania	ETH(FB)	1959	1970	1979	1989	2001
Macedonia	FB			1981*	1994	
Moldova	ETH(FB)	1959	1970	1979	1989	
Poland	FB		1970			2002
Romania	FB		1966		1992	2002
Russian Federation	ETH(FB)	1959	1970	1979	1989	2002
Serbia & Montenegro	FB			1981*	1991	2002
Slovakia	FB				1991*	2001
Slovenia	FB			1981*	1991	2002
Tajikistan	ETH(FB)	1959	1970	1979	1989	
Turkey	FB	1960	1965	1980	1990	2000
Turkmenistan	ETH(FB)	1959	1970	1979	1989	
Ukraine	ETH(FB)	1959	1970	1979	1989	2001
Uzbekistan	ETH(FB)	1959	1970	1979	1989	
High income Middle East and North Africa						
Bahrain	NT	1959	1971	1981	1991	2001
Israel	FB	1961	1972	1983		2001
Kuwait	NT	1957	1970	1975	1985	2001
Oman	NT				1993	2004
Qatar	FB					
Saudi Arabia	NT				1992	1995
United Arab Emirates	NT			1980	1993	2003
Rest of Middle East and North Africa						
Algeria	NT		1966			
Egypt	NT	1960		1976	1986	1996
Iran (Islamic Republic of)	NT				1986	1996

Iraq	FB	1957				1997
Jordan	NT	1961		1979	1994	2004
Lebanon	FB					1996
Libyan Arab Jamahiriya	NT	1964	1973			
Morocco	NT	1960	1971			2004
Occupied Palestinian Territory	FB					1997
Syrian Arab Republic	NT	1960	1970	1981	1994	
Tunisia	NT	1956	1966	1984	1994	2004
Yemen	NT				1986	2004

Africa						
Angola	FB	1960		1983	1993	
Benin	NT			1979		2002
Botswana	NT		1971	1981	1991	2001
Burkina Faso	FB			1975	1985	1996
Burundi	FB			1979	1990	
Cameroon	FB			1976	1987	
Cape Verde	NT			1980	1990	
Central African Republic	NT			1975	1988	
Chad	FB				1993	
Comoros	FB	1958		1980	1991	
Congo	NT		1974	1984		
Côte d'Ivoire	NT			1975	1988	1998
Democratic Republic of the Congo	NT	1958*		1984		
Djibouti	FB				1991	
Equatorial Guinea	NT	1950		1983		
Eritrea	FB					
Ethiopia	NT	1961			1994	

Gabon	NT	1960			1993	
Gambia	NT	1963	1973	1983	1993	
Ghana	FB	1960	1970	1984		2000
Guinea	NT			1983		1996
Guinea-Bissau	FB	1950		1979	1991	
Kenya	FB	1962	1969	1979	1989	1999
Lesotho	NT	1956		1976	1986	1996
Liberia	FB	1962	1974	1984		
Madagascar	NT		1965	1975	1993	
Malawi	FB		1966	1977		
Mali	FB			1976	1987	1998
Mauritania	NT			1977	1988	
Mauritius	NT		1972	1983	1990	2000
Mayotte	FB				1991	1997
Mozambique	NT	1955		1980		1997
Namibia	NT				1991	2001
Niger	NT			1977	1993	2001
Nigeria	NT	1963			1991	
Rwanda	NT	1958*		1978	1991	2002
Réunion	FB	1961	1974	1982	1990	1999
Saint Helena	FB		1966	1976	1987	1998
Sao Tome and Principe	NT			1981	1991	
Senegal	FB	1960		1976	1988	2002
Seychelles	NT	1960		1982	1987	1997
Sierra Leone	FB				1985	2004
Somalia	FB					
South Africa	FB	1961	1970	1980	1985	2001

Sudan	FB	1956		1983	1993	
Swaziland	FB	1956	1966	1976	1986	1997
Togo	NT			1981		
Uganda	NT		1969		1991	2002
United Republic of Tanzania	FB		1967	1978	1988	2002
Zambia	FB	1963	1969	1980	1990	
Zimbabwe	FB	1956	1969		1992	
South Asia						
Afghanistan	FB			1975		
Bangladesh	FB	1961	1974			
Bhutan	FB					2005
India	FB	1961	1971	1981	1991	2001
Maldives	FB					
Nepal	FB	1961	1971	1981	1991	2001
Pakistan	FB	1961	1973			1998
Sri Lanka	NT	1963	1971	1981		
East Asia and the Pacific						
American Samoa	FB	1960	1970	1980	1990	2000
Brunei Darussalam	FB	1960	1971	1981	1991	
Cambodia	FB					1998
China	FB					
China, Hong Kong Special Administrative Region	FB	1961	1971	1981	1991	2001
China, Macao Special Administrative Region	FB			1981	1991	2001
Cook Islands	FB	1956	1966	1976		1996
Democratic People's Republic of Korea	FB					

Democratic Republic of Timor-Leste	FB					2004
Fiji	FB	1956	1966	1976	1986	
French Polynesia	FB	1962		1977	1988	1996
Guam	FB	1960	1970	1980	1990	2000
Indonesia	NT		1971		1990	2000
Kiribati	FB	1963	1973	1978	1990	2000
Lao People's Democratic Republic	NT					1995
Malaysia	FB	1957	1970	1980	1991	2000
Marshall Islands	NT				1988	1999
Micronesia (Federated States of)	FB		1973		1994	2000
Mongolia	NT					2000
Myanmar	NT		1973		1994	2002
Nauru	FB	1961	1966	1977		2002
New Caledonia	FB	1963	1969	1983	1989	1996
Niue	FB	1956	1966	1976	1986	
Norfolk Island	FB			1981	1991	2001
Northern Mariana Islands	FB			1980	1990	2000
Palau	FB			1980	1990	2000
Papua New Guinea	FB		1966	1980		
Philippines	NT	1960	1970	1980	1990	2000
Republic of Korea	NT(FB)	1960	1970	1980	1990	2000
Samoa	FB	1956	1971		1986	2001
Singapore	FB	1957	1970	1980	1990	2000
Solomon Islands	FB		1970	1976	1986	1999
Taiwan	NT				1990	2000
Thailand	NT	1960	1970			2000
Tokelau	FB	1961	1972	1976	1986	2001

Tonga	FB	1956	1966	1976	1986	1996
Tuvalu	FB	1963*	1973*			
Vanuatu	FB		1967	1979	1989	1999
Viet Nam	FB				1989	
Wallis and Futuna Islands	FB		1969	1976	1990	2003
Latin America and the Caribbean						
Anguilla	FB			1984	1992	2001
Antigua and Barbuda	FB	1960	1970		1991	2001
Argentina	FB	1960	1970	1980	1991	2001
Aruba	FB	1960		1981	1991	2000
Bahamas	FB	1960	1970	1980	1990	
Barbados	FB	1960		1980	1990	
Belize	FB	1960		1980	1991	2000
Bermuda	FB	1960	1970	1980	1991	2000
Bolivia	FB	1950		1976	1992	2001
Brazil	FB	1960	1970	1980	1991	2000
British Virgin Islands	FB	1960	1970	1980	1991	
Cayman Islands	FB	1960		1979	1989	2000
Chile	FB	1960	1970	1982	1992	2002
Colombia	FB	1964	1970		1993	2005
Costa Rica	FB	1963	1973	1984		2000
Cuba	FB	1953	1970			2000
Dominica	FB	1960		1981	1991	
Dominican Republic	FB	1960	1970			2002
Ecuador	FB	1962	1974	1982	1990	2001
El Salvador	FB	1961	1971		1992	2007
Falkland Islands (Malvinas)	FB	1962	1972		1986	2001

French Guiana	FB	1961	1974	1982	1990	1999
Greenland	FB	1960	1970	1976		
Grenada	FB	1960		1981	1991	
Guadeloupe	FB	1961	1974	1982	1990	1999
Guatemala	FB	1963	1973	1981	1994	2002
Guyana	FB	1960		1980	1991	2002
Haiti	FB	1950	1971	1982		
Honduras	FB	1961			1988	2001
Jamaica	FB	1960	1970	1982	1991	2001
Martinique	FB	1961	1974	1982	1990	1999
Mexico	FB	1960	1970	1980	1990	2000
Montserrat	FB	1960	1970	1980	1991	
Netherlands Antilles	FB		1971	1981	1992	2001
Nicaragua	FB	1963	1971			1995
Panama	FB	1960	1970	1980	1990	2000
Paraguay	FB	1950	1972	1982	1992	2002
Peru	FB	1960	1972	1981	1993	
Puerto Rico	FB		1970	1980	1990	2000
Saint Kitts and Nevis	FB	1960	1970	1980	1991	2001
Saint Lucia	FB	1960		1980	1991	2001
Saint Pierre et Miquelon	FB	1962	1974	1982		1999
Saint Vincent and the Grenadines	FB	1960		1980	1991	
Suriname	NT	1964				2004
Trinidad and Tobago	FB	1960	1970	1980	1990	2000
Turks and Caicos Islands	FB	1960		1980	1990	
United States Virgin Islands	FB	1960	1970	1980	1990	2000
Uruguay	FB	1963		1975	1985	1996

Venezuela	FB	1961	1971	1981	1990	2001	

*The census year was derived from splitting an aggregated census.
a. FB is foreign born, NT is nationality, and ETH is ethnic group . FB(NT) means that the original data by nationality were amended and the resulting numbers are closer to foreign-born definition.
Source: Authors' calculations based on data described in text.

TABLE A2. List of Aggregations

Aggregated region	Master region	Aggregated region	Master region
Aden	Yemen	**Northern Ireland**	United Kingdom
Alaska	United States of America	**Palmyra**	United States of America
Alboran and Perejil	Spain	**Panama Canal Zone**	Panama
Ascension Island	Saint Helena	**Penang**	Malaysia
Azores	Portugal	**Pitcairn Island**	United Kingdom
Bonaire	Netherlands Antilles	**Providencia Island**	Colombia
Born abroad of U.S. parent(s)	United States of America	**Saint Croix**	United States Virgin Islands
British Indian Ocean Territory	United Kingdom	**Saint Martin**	Netherlands Antilles
Canary Islands	Spain	**Saint Thomas**	United States Virgin Islands
Canton and Enderbury Islands	Kiribati	**San Andres Island**	Saint Pierre and Miquelon
Ceuta and/or Melilla	Spain	**Sarawak**	Malaysia
Channel Islands	United Kingdom	**Scotland**	United Kingdom
Channel Islands and the Isle of Man	United Kingdom	**South Senegal**	Senegal
Christmas Island	Australia	**South Vietnam**	Vietnam
Cocos (Keeling) Islands	Australia	**South Yemen**	Yemen
Curacao	Netherlands Antilles	**Spanish Sahara**	Morocco
Dubai	United Arab Emirates	**Svalbard and J. Mayen Islands**	Norway
East Germany	Germany	**Terre Nova**	Canada
Easter Island	Chile	**Tristan de Cunha**	Saint Helena
England	United Kingdom	**Vatican**	Italy
England and Wales	United Kingdom	**Wake Island**	United States of America
French India	India	**Wales**	United Kingdom
Galapagos	Ecuador	**West Germany**	Germany
Gaza Strip	Occupied Palestinian Territory	**Western New Guinea**	Indonesia
Germany (East Berlin)	Germany	**Western Sahara**	Morocco
Germany (unspecified)	Germany	**Zanzibar**	Tanzania
Great Britain	United Kingdom		
Hawaii	United States of America		
Howland Island	United States of America		
Isle of Man	United Kingdom		
Jammu	India		
Johnston Islands	United States of America		
Kashmir	India		
Kosovo	Serbia and Montenegro		
Labuan	Malaysia		
Madeira	Portugal		

North Borneo	Malaysia
North Senegal	Senegal
North Vietnam	Vietnam
North Yemen	Yemen

Source: Authors' calculations based on data described in text.

Table A3. Allocation of Aggregate Origin Region by Migrant Shares over Time for Australia

Origin country listed in 2001 Australian census	Total immigrants to Australia in 2001	Share of Soviet Union migration to Australia in 2001 (%)	Number of migrants allocated in 1966 across constituent countries
Azerbaijan	145	0.3	93
Armenia	899	2.0	576
Belarus	1,041	2.3	667
Estonia	2,386	5.2	1,529
Georgia	310	0.7	199
Kazakhstan	438	1.0	281
Kyrgyzstan	101	0.2	65
Latvia	6,690	14.6	4,287
Lithuania	3,689	8.1	2,364
Moldova	483	1.1	309
Russian Federation	15,022	32.8	9,625
Tajikistan	41	0.1	26
Turkmenistan	26	0.1	17
Ukraine	14,062	30.7	9,010
Uzbekistan	412	0.9	264
Total USSR	**45,745**	**100**	**29,311**

Source: Authors' calculations based on data described in text.

Table A4. Calculation of Sex Ratios Based on Concurrent Subregional Shares.

Destination country in Scandinavia	Number of male migrants in 1990 from Uruguay	Number of female migrants in 1990 from Uruguay	Males (%)	Females (%)
Denmark	92	90	51	49
Finland	11	21	39	66
Norway	67	78	46	54
Average across subregion			**47**	**53**

Source: Authors' calculations based on data described in text.